**CS416: Data Visualization**

**Summer 2021**

**Final Report – Narrative Project Essay**

**“Surviving the Titanic”**

A large ship in the water

Description automatically generated with medium confidence[[1]](#footnote-1)Graphical user interface, application, Word

Description automatically generated

Theodore Hromadka / [tvh3@illinois.edu](mailto:tvh3@illinois.edu) / **https://hromadka.github.io/**

# Introduction

This essay describes the seven aspects of the data visualization narrative prepared as the final project for this course. This project presents a drill-down narrative about the survivors of the sinking of the RMS Titanic in 1912.

The visualization is available at [**https://hromadka.github.io/**](https://hromadka.github.io/)

The chosen dataset was created expressly for this course, as existing datasets did not include the crew data. For example, the Titanic dataset on Kaggle is a famous introductory example to machine learning. However, this dataset only includes labels for a portion of the passengers, and none of the crew. Sources for our custom data follow in Appendix A.

All values presented in this visualization are calculated on the fly using D3.JS rollups and filters reading the full 2,169-person dataset, with one record per person. Charts are not based on simple, hard-coded maps or pre-calculated values. The charts are completely dynamic and represent several days of fighting temptation.

# Messaging

This visualization attempts to communicate the different situational factors affecting the odds for an arbitrary person aboard the Titanic at the time it struck the iceberg surviving until being rescued by the RMS Carpathia[[2]](#footnote-2).

We consider three main factors: age, gender, and status (passenger ticket class or working crew).

Our main message is that this data confirms the widely-held belief that odds of survival increased for higher-class and/or female passengers. Also, it turned out that age was not a very significant factor in survival.

# Narrative Structure

We implemented a drill-down narrative structure in order to allow the user to explore different individual factors, or combinations of those factors. We start with an initial landing page that summarizes the project and presents a simple pie chart. This pie chart displays overall survival rates for the entire ship’s manifest. Clicking on that pie chart, as indicated, leads the user to a more detailed bar chart of survival odds, with options available for re-projecting the data based on a chosen dimension (Age, Gender, Class). In the case of Gender, clicking on the re-projected bar chart allows the user to drill down even further on another page of bar charts, exploring the data with the combinations of (Gender, Age) or (Gender, Class).

# Visual Structure

The visual structure for each scene follows a basic template: drill-down options at the top, explanatory text and instructions below that, and then a large central workspace displaying a chart type appropriate for the data being displayed. Since the user reads top-down, placing the instructions at the top provides immediate guidance as to how the page should be used. The pages are kept simple and clean. Axes are labeled and axis tick marks are evenly spaced in legible font sizes. Survival percentage values are presented as annotations at the end of each chart bar. User attention is then focused on the chart, which dominates the scene and highlights what the user should focus on.

Mouse hover events perform two actions: the selected chart bar changes color, and a tooltip is displayed providing additional data (exact person counts) next to the selected “survival percentage” chart bar.

Chart bar colors are kept consistent throughout, and further refined on the charts that combine factors to indicate distinct data categories (e.g., male data is colored blue, and female data is colored pink). The transition from one scene to the next is eased by presenting the same level of depth of the previous scene on the initial display of the more complicated successive scene, i.e. the second scene starts with the same overall view percentages of the first scene, and the third scene starts with the same single-factor percentages of the second scene. Explicit instructions on where to click help the user navigate to the other scenes.

# Scenes

Three scenes are included in this presentation. The scenes are ordered in sequence of: (1) **overall** data, (2) specific **single factor** data, and (3) specific **combinations of two factors** data. This follows an explicit ordering of the drill-down process, with each scene showing larger amounts of chart data as the information dive goes deeper.

Graphical user interface, text, application

Description automatically generated Graphical user interface

Description automatically generated Chart

Description automatically generated

# Annotations

The annotation template is that the percentage values are displayed on the pie chart wedge (scene 1) or at the end of each horizontal bar on the bar charts (scenes 2 and 3). These annotations support the messaging in that the main message of this presentation is the survival odds of a passenger, refined by whichever factor(s) is/are chosen by the user.

The annotation values can change within a single scene, because the user is allowed to “drill down” on that data by choosing refinement factors. For example, selecting the “Age” option transitions the bar chart into a display of survival odds for each 10-year binned age segment. In that case, the percentage values of the prior display would transition to the Age survival odds values.

# Parameters

The parameters of the visualization are the selected option(s) for drilling down into the data. The states of the narrative are permutations of the three factors: Age, Gender, and Class. The parameters are controlled by the option buttons presented at the top of the screen. These parameters are passed to the renderData() function as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| **Age option state** | **Gender option state** | **Class option state** | **States[] parameters** |
| Unselected (default) | Unselected (default) | Unselected (default) | [ 0 0 0 ] |
| **SELECTED** | Unselected | Unselected | [ 1 0 0 ] |
| Unselected | **SELECTED** | Unselected | [ 0 1 0 ] |
| Unselected | Unselected | **SELECTED** | [ 0 0 1 ] |
| **SELECTED** | **SELECTED** | Unselected | [ 1 1 0 ] |
| Unselected | **SELECTED** | **SELECTED** | [ 0 1 1 ] |
| … | … | … | … |

The default state (all unselected) indicates that the “overall” high-level view should be presented. A state with a single “1” indicates which single dimension should be displayed in Scene 2. A state with two “1” values indicates which two dimensions should be combined and displayed in Scene 3.

# Triggers

The main triggers of control are the selection of a given radio option button(s). This leads to the chart transitioning to the new survival percentage values based on the chosen dimension or combination of dimensions.

Affordances are most obviously the option buttons themselves, along with the “Please select a feature” caption next to those buttons. We also include an explicit “Please highlight a chart bar for more information” caption in order to guide the user towards the “mouseOver” events, which would present additional detailed information about the data “under” the mouse cursor at that moment. Lastly, we also present captions indicating that the user can click on the chart in order to “drill down” by moving to the following, more detailed scene.

1. Public domain photo of RMS Titanic, By F.G.O. Stuart (1843-1923) - http://www.uwants.com/viewthread.php?tid=3817223&amp;extra=page%3D1, Public Domain, https://commons.wikimedia.org/w/index.php?curid=2990792 [↑](#footnote-ref-1)
2. https://en.wikipedia.org/wiki/Titanic [↑](#footnote-ref-2)